XXVI. "On the Relation between the Electrical Resistances of Pure Metals and their Molecular Constants." By W. WILLIAMS. Communicated by Professor ANDREW GRAY, F.R.S.

The Society adjourned over the Long Vacation to Thursday, November 21, 1901.

"On the Mathematical Theory of Errors of Judgment, with Special Reference to the Personal Equation." By KARL PEARSON, F.R.S., University College, London. Received April 23,—Read June 20, 1901.

(Abstract.)

In 1896 I, with Dr. Alice Lee and Mr. G. A. Yule, made a series of experiments on the bisection of lines at sight. The object of these experiments was to test a development of the current theory of errors of observation, by which it seemed possible to me to determine the *absolute* steadiness of judgment of any individual by comparing the relative observations of three (instead of as usual two) observers. As a rule the absolute error of the observer is unknown and unknowable, and I was seeking for a quantitative test of steadiness in judgment to be based on relative judgments. If σ_{01} be the standard deviation of the absolute judgments of the first observer, σ_{12}, σ_{23}, σ_{31} the standard deviations of the relative judgments of the first and second, the second and third, and the third and first observers respectively, then

$$\sigma_{01}^2 = \tfrac{1}{2}(\sigma_{21}^2 + \sigma_{13}^2 - \sigma_{23}^2) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad \text{(i)}$$

on the basis of the current theory of errors. Thus it seemed possible to determine absolute steadiness of judgment from the standard deviations of *relative* judgments, which are all that the physicist or astronomer can usually make, provided three observers and not two were compared.

To my great surprise I found results such as (i) were not even approximately true, and that they failed to hold because the judgments of the observers were *substantially correlated*. It did not occur to me at first that judgments made as to the midpoints of lines by experimenters, in the same room it is true, but not necessarily bisecting the same line at the same instant, could be psychologically correlated, and I looked about for a source of correlation in the treatment of the data. We had taken 500 lines of different lengths and bisected them at sight; assuming that the error would be more or less proportional to the length of the line, I had adopted the deviation from the

true midpoint to the right in terms of the length of the line as the error. I was then led to realise the importance of what I have termed "spurious correlation" in this use of indices or ratios, and I published a short notice of the subject in the 'Roy. Soc. Proc.,' vol. 60, p. 489, 1896.

It seemed necessary accordingly to make our judgments in a different manner, and a second series of 520 experiments was made by Dr. Alice Lee, Dr. W. F. Macdonell, and myself, in which we observed the motion of a narrow beam of light down a uniform strip of fixed length, and recorded its position at the instant, *à priori* unknown to us, at which a hammer struck a small bell. The experiment was made by means of a pendulum devised by Mr. Horace Darwin, and the record required a combination of ear, eye, and hand judgment. In the manipulation of the data there was no room for the appearance of "spurious correlation," but to my great surprise I again found substantial correlation in two out of the three cases of what one might reasonably suppose to be absolutely independent judgments.

This led to a thorough reinvestigation of the bisection experiments, absolute and not ratio errors being now dealt with. We found the same result, *i.e.*, correlation of apparently independent judgments. The absolute personal equations based on the average of twenty-five to thirty experimental sets were then plotted, and found to fluctuate in sympathy, and these fluctuations were themselves far beyond the order of the probable errors of random sampling. Nor were the fluctuations explicable solely by likeness of environment. For in the bright line experiments while the judgments of A and B were sensibly uncorrelated, those of C were substantially correlated with those of both A and B. Thus we were forced to the conclusion that judgment depends in the main upon some few rather than upon many personal characteristics, and that while A and B had practically no common characteristics, there were some common to A and C and others common to B and C. We are driven to infer—

(i.) That the fluctuations in personal equation are not of the order of the probable deviations due to random sampling.

(ii.) That these fluctuations in the case of different observers, recording absolutely independently, are sympathetic, being due to the influence of the immediate atmosphere of the observation or experiment on personal characteristics, probably few in number, one or more of which may be common to each pair of observers.

In this way we grasp how the judgments of "independent" observers may be found to be substantially correlated. In the memoir attention is drawn to the great importance of this, not only for the weighting of combined observations, but also for the problem of the stress to be laid on the testimony of apparently independent witnesses to the same phenomenon.

The current theory of the personal equation thus appears to need modification, and we require for the true consideration of relative judgments not only a knowledge of the variability of observers, but also of their correlation in judgment as necessary supplements to the simple personal equation.

Having obtained from our data twelve series of errors of observation considerably longer than those often or even exceptionally dealt with by observers, we had a good opportunity for testing the applicability of the current theory of errors, in particular the fitness of the Gaussian curve

$$y = y_0 \, e^{-x^2/(2\sigma^2)}$$

to describe the frequency of errors of observation. In a considerable proportion of the cases this curve was found to be quite inapplicable. Errors in excess and defect of equal magnitude were not equally frequent; skewness of distribution, sensible deviation of the mode from the mean, "crowding round the mean," even in the case of passable symmetry, all existed to such an extent as to make the odds against the error distributions being random samples from material following the Gaussian law of distribution enormous. It is clear that deviation of the mode from the mean, and the independence of at least the first four error moments, must be features of any theory which endeavours to describe the frequency of errors of observation or of judgment within the limits allowable by the theory of random sampling. The results reached will serve to still further emphasise the conclusions I have before expressed:

(*a*.) That the current theory of errors has been based too exclusively on mathematical axioms, and not tested sufficiently at each stage by comparison with actual observations or experiments.

(*b*.) That the authority of great names—Gauss, Laplace, Poisson—has given it an almost sacrosanct character, so that we find it in current use by physicists, astronomers, and writers on the kinetic theory of gases, often without a question as to its fitness to represent all sorts of observations (and even insensible phenomena!) with a high degree of accuracy.

(*c*.) That the fundamental requisites of an extended theory are that it must—

(i.) Start from the three basal axioms of the Gaussian theory and enlarge and widen them.

(ii.) Provide a systematic method of fitting theoretical frequencies to observed distributions with (α) as few constants as possible, (β) these constants easily determinable and closely related to the physical characters of the distribution, and

(iii.) When improbable isolated observations are rejected, give theoretical frequencies not differing from the observed frequencies by more than the probable deviations due to random sampling.

I propose to consider these points in reference to the skew frequency distributions discussed in a memoir in the 'Phil. Trans.' for 1895 (A, vol. 186, *et seq.*) in another place. The present memoir, however, shows that these skew distributions give results immensely more probable than the Gaussian curve, and thus confirms in the case of errors of observation the results already reached in the case of organic variation.

"Mathematical Contributions to the Theory of Evolution.—X. Supplement to a Memoir on Skew Variation." By KARL PEARSON, F.R.S., University College, London. Received May 22,—Read June 20, 1901.

(Abstract.)

In the second memoir of this series a system of curves suitable for describing skew distributions of frequency was deduced from the solutions of the differential equation

$$\frac{1}{y}\frac{dy}{dx} = \frac{b_0 + b_1 x}{a_0 + a_1 x + a_2 x^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}.$$

These solutions were found to cover satisfactorily a very wide range of frequency distributions of all degrees of skewness. Two forms of solution of this differential equation, depending upon certain relations among its constants, had, however, escaped observation, for the simple reason that all the distributions of actual frequency I had at that time met with fell into one or other of the four types dealt with in that memoir. A little later the investigation of frequency in various cases of botanical variation showed that none of the four types were suitable, and led me to the discovery that I had not found all the possible solutions of the differential equation above given. Two new types were found to exist—

Type V: $$y = y_0 x^p e^{-\gamma/x} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)},$$

with a range from $x = 0$ to $x = \infty$, and

Type VI: $$y = y_0 (x-a)^{m_1} x^{-m_2} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(iii)},$$

with a range from $x = a$ to $x = \infty$.

These curves were found to be exactly those required in the cases which my co-workers and I in England, and one or two biologists in America, had discovered led in the earlier Types I and IV to impossible results, *i.e.*, to imaginary values of the constants.

In the present memoir the six types are arranged in their natural order, and a criterion given for distinguishing between them. They are illustrated by three examples: (a) age of bride on marriage for a